DOCUMENT RESUME

ED 218 898 EC 143 163

AUTHOR Lichtenstein, Robert
TITLE Validity of Preschool Screening Procedures: Research with the Minneapolis Preschool Screening Instrument.
PUB DATE Apr 82
NOTE 32p.; Paper presented at the Annual International Convention of the Council for Exceptional Children (60th, Houston, TX, April 11-16, 1982, Session W-21). The MPSI (Minneapolis Preschool Screening Instrument) the author refers to in this paper is available through the Minneapolis Public Schools/Prescriptive Instruction Center, 254 Upton Ave., South, Minneapolis, MN 55455 (MPSI test kit $35.00; MPSI Response forms $3.50 package of 30).

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Disabilities; Preschool Education; *Screening Tests; *Test Construction; Test Reliability; Test Validity
IDENTIFIERS *Minneapolis Preschool Screening Instrument

ABSTRACT

The paper describes development of the Minneapolis Preschool Screening Instrument (MPSI), a measure for identifying children needing preschool intervention. The measure was designed to be brief to administer, inexpensive and simple enough so that both professionals and nonprofessionals could quickly learn administration and scoring. A psychoeducational focus was taken in devising developmental tasks to measure cognitive, fine motor, language, memory, and perceptual functioning. Eleven subtests were designed: Building, Copying Shapes, Information, Matching, Sentence, Completion, Hopping and Balancing, Naming Colors, Counting, Prepositions, Identifying Body Parts, and Repeating Sentences. Research to test MPSI reliability and validity included comparisons with Developmental Indicators for the Assessment of Learning. A followup study revealed the MPSI failed to identify some children who were rated as having moderate to severe problems in kindergarten and many children with only mild problems. (SB)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

This document has been reproduced as received from the person or organization originating it.
Minor changes have been made to improve reproduction quality.

Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

ED218898

VALIDITY OF PRESCHOOL SCREENING PROCEDURES: RESEARCH WITH THE MINNEAPOLIS PRESCHOOL SCREENING INSTRUMENT

Robert Lichtenstein

University of Delaware

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Robert Lichtenstein

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Paper presented at the 60th Annual International Conference of the Council for Exceptional Children, Houston, April 1982.

EC143163

This presentation involves my long-term program of research with the Minneapolis Preschool Screening Instrument (Lichtenstein, 1980a), which was developed to address the concern that existing measures were not adequate for the task of screening preschool children for psychoeducational problems. I will take this up chronologically to give you a sense for how the MPSI and my views of preschool screening developed simultaneously. It is best to begin this story at the beginning, rather than start in with my philosophical conclusions and skeptical comments and endless cautions. That is the last thing you want to hear when you are assigned the responsibility of carrying out the mandate (i.e., of Public Law 94-142) to identify all preschool children with special educational needs . . . and to do so in an economical, non-discriminatory, reliable, and valid manner. And this is where I started as well--as a professional dealing with such questions as which children to serve, what screening measures to use, and how to go about making screening decisions. (Actually, I'm going to begin this story a bit earlier than that even, to provide a brief illustration of how public policy and government funding patterns shape careers and focus research interests.)

During the first year that Public Law 94-142 (Education for All Handicapped Children Act) was in effect, I was on an "indefinite leave of absence" from school psychology and from graduate school. This non-stipended sabbatical came to an end as I accepted a research position for Project SEARCH, a research project administered by the Minneapolis Public Schools which was funded by a grant from Bureau of Education for the Handicapped (BEH) as part of an effort to help implement new special education policies. This project's assigned task was to develop an early identification system for Minneapolis that would also serve as a resource for programs throughout the state. One Project SEARCH activity, compiling a directory of screening instruments

(Davidson, Lichtenstein, Canter, & Cronin, 1977), led to the striking conclusion that existing instruments were unsatisfactory for meeting the particular needs of our school district and of many others in Minnesota (the latter judging from the comments of other despairing practitioners in the state who contacted us in the hope that we had been more successful than they in this task). School districts in Minnesota were selecting screening measures on the basis of testimonial evidence (i.e., from sharing notes with the next district over), nebulous research data cited in publishers' brochures, process of elimination, and other random strategies that people rely on when presented with insufficient information and a lack of good choices. Some districts even resolved to develop their own screening measures, since they knew what they wanted and it did not exist. Through some peculiar twists of fate, the Minneapolis Public Schools decided to pursue this route and I was assigned primary responsibility for developing what was to eventually become the Minneapolis Preschool Screening Instrument (MPSI).

Test Construction Strategy

I began by identifying practical, theoretical, and empirical characteristics of the type of preschool screening instrument that was desired. Practical considerations revolved around the fact that screening is largely an economic matter--we screen because it is impractical and unfeasible to comprehensively assess every child. The objective is to obtain high accuracy in making screening decisions with a minimum of time and expense. Therefore, the first guidelines were that the measure should be (1) brief to administer--certainly no longer than 20 minutes, (2) inexpensive--in terms of equipment and materials, consumables, and personnel, and (3) simple--so that both professionals and non-professionals could learn it quickly and administer and score it flawlessly. Another consideration involved the age range to be

encompassed by a preschool screening instrument. The range of approximately 3-3/4 to 5-1/4 years of age was defined on the basis of two considerations: the child should be screened far enough in advance of the time of school entry to permit preschool intervention to take place, but not so early that psychoeducational material was inappropriate and invalid.

One strong bias in my approach, evolving from my school psychology background and from my knowledge of measurement research, was towards direct assessment using standardized measures. By "standardized," I mean two things: (1) assessment of each child is based on his/her performance under the same stimulus conditions, using uniform evaluation criteria, and (2) results are interpreted by comparing a child's performance with other children the same age in a normative sample. Although there is considerable controversy over standardized testing, it is widely recognized that educational measurement is a rather primitive art without objective standards for comparison.

Another key consideration was the type of content desired for a preschool instrument with a psychoeducational focus. Appropriate for inclusion in such a measure are those developmental tasks which tap the full range of skills and abilities that are critical to successful school performance--a multidimensional array that includes cognitive, fine motor, language, memory, and perceptual functioning. Closely related to content is the matter of appropriate difficulty level. A screening test should be designed to make discriminations at the low end of the continuum, that is, to distinguish between children functioning well below age expectations and those functioning within normal expectations. This is accomplished by selecting items which the majority of children in the population will pass. Since the

measure was intended for use with a narrow age range, it was possible to apply the guideline of selecting items that were relatively easy for the target population, that is, items that would be passed by a high proportion of four year olds.

A major inconsistency among various screening instruments involves the type of scoring system used. The particular question is whether items should be divided into separate scales (e.g., language, fine motor, gross motor), with each scale separately scored, or whether a single overall score should be used. Studies with young children indicate that ability factors most directly related to academic achievement--reasoning, language, memory, perceptual/performance, quantitative--are both intercorrelated and are independent to some degree (Kaufman, 1975; Meyers, Orpet, Attwell, & Dingman, 1964; Ramsey & Vane, 1970; Wallbrown, Blaha, & Wherry, 1973). Therefore, the case could be made to use either multiple scales or a single unidimensional scale. Two considerations guided my decision to devise a single overall scale: (1) longer scales have higher reliability, and (2) the information from multiple scales would eventually be combined into a single overall screening decision (i.e., to refer or not refer). To break this set into more specific components might be of interest for diagnostic purposes, but not appropriate for screening. (Let me, however, make the distinction that there are other areas of functioning which are less related to this set and could not be encompassed within this scale, e.g., social skills, emotional adjustment, gross motor skills, and speech articulation.)

The steps involved in the actual selection of test content and formulation of administration and scoring procedures were rather straightforward, but of course critical. As will soon be evident, there was nothing new or different or exceptional about the instrument that was being designed. The

intention from the start was to apply the cumulative wisdom of decades of educational and psychological research to construct an instrument ideally suited for the purpose of screening preschool children . . . and to do so without repeating past mistakes or inventing new ones. This, if successfully done, was expected to produce a measure that worked better than any previous measure for its designated purpose.

The wisdom of the ages strongly suggested an empirical approach to the selection and evaluation of test content. First, a large pool of promising developmental tasks was identified which met the previously-stated criteria. Data from published research and from original research studies in local screening programs were then used to select and develop optimal items in terms of (1) reliability, (2) validity, (3) difficulty level, and (4) overall coverage of target content areas. As examples, counting tasks were shown to be consistently good predictors of later achievement; questions such as "What do we do with our ears?" proved to be best representative of Stanford-Binet verbal reasoning items passed by most four year olds (Ramsey & Vane, 1970); a validity study of a local school district's homemade screening test indicated that a sentence repetition task was an excellent predictor of kindergarten teacher ratings of language skills. An excess of items was retained at this point with the intention that only the "fittest" would survive the following stages.

Extensive and careful pilot testing is a must for avoiding both familiar and newfound pitfalls--such as instructions to children that are so stilted that examiners feel compelled to modify them; such as stimulus pictures for a vocabulary subtest that could just as easily be a cat, a rat, or even a bat; such as subtests that make impossible demands on examiners (demonstrating the hop-skip-and-jump while reading involved instructions and recording where

the child lands). The "debugging" process is critical for insuring quality control. The MPSI was pilot tested and revised repeatedly, undergoing a metamorphosis that involved six versions in all. Input from special education teachers and administrators, psychologists, speech and language specialists, measurement experts, and paraprofessional examiners resulted in continual upgrading of test content, recording procedures, standardized instructions, scoring criteria, materials, etc. Some examples: a four-page scoring form was reduced to a single sheet for economy and simplicity of handling; a draw-a-person subtest was found to be too time-consuming and too difficult to reliably score; additional scoring rules were added to cover new, unusual, and creative responses provided by examinees. Early in the process, one reviewer made the observation that the data-based approach to selecting test content resulted in a set of subtests that "didn't look like much fun." She was right--this "all-star lineup" of tasks which worked well in other situations was psychometrically powerful, but unlikely to hold a young child's interest. More high interest tasks were added to address this concern.

Pilot testing also provided more item analysis data for weeding out undesirable items. As examples: (1) having the child point to his/her nose was too easy, (2) sentence repetition items were more discriminating (i.e., had higher item validity figures) than number repetition items, and (3) naming the "colors" white and black worked better than blue and red.

As for the matter of determining whether the resulting instrument succeeded as intended, an ambitious research program was planned, including a test-retest and inter-rater reliability study, several concurrent validity studies, and a large-scale predictive validity study. But first, let me describe the MPSI.

The Minneapolis Preschool Screening Instrument

The current version of the MPSI, as shown by the scoring form in Figure 1, contains 50 dichotomously-scored items (pass = 1, fail = 0) distributed over 11 subtests: Building, Copying Shapes, Information, Matching, Sentence Completion, Hopping and Balancing, Naming Colors, Counting, Prepositions, Identifying Body Parts, and Repeating Sentences. Scores on these items are totaled, and a screening result of either pass or refer is determined from the cutoff score for the child's age range. In addition, a speech intelligibility rating (on a scale of 1 to 4) is assigned by the examiner. Administration time for the complete instrument is approximately 12-15 minutes.

The MPSI is designed so that it can be reliably administered by paraprofessionals who have completed a four-hour training session and a one-hour follow-up session. Procedures for training examiners are outlined in detail in the manual. This insures some consistency in training, and also enables examiners to refer back to the training guidelines for review.

Normative data for the MPSI were collected on 1320 Minneapolis preschool children screened over two consecutive school years (see Table 1). The sample consisted of children ages 3-7 to 5-4 from a cross-section of socioeconomic levels from lower to upper-middle class. Approximate breakdown of the racial composition of this group was 86% white and 14% black and other minorities. Cutoff points, established for each of seven age ranges, were set to refer approximately 15% of the children in the Minneapolis normative sample.

Administration of the MPSI begins with the examiner interacting with the child to establish rapport. The initial subtests--Building and Copying Shapes--further contribute to building of rapport. They involve fine motor responses

rather than verbal responses, which helps the shy or inhibited child become more comfortable in the testing situation. In an early version of the MPSI, the Hopping and Balancing subtest was placed at the beginning. This proved to be a gross mistake; for many preschoolers, gross motor tasks are the most threatening tasks in the measure.

The Building subtest has the child construct wooden block towers to match the sample built by the examiner. Copying Shapes is also a conceptual/fine motor matching task, with the child drawing figures to match those printed on the reverse of the scoring form. Copying Shapes items, originally scored 0/1/2, were simplified to 0/1 scoring to insure high reliability. Extensive sets of critical scoring samples (i.e., "just passing" and "just failing" responses) were compiled to supplement the scoring criteria to facilitate scoring.

Item analysis data indicate that the general knowledge questions of the Information subtest are among the strongest items. A late change was made in item 12 of this subtest. Screening program examiners were noting that the question "Which is bigger, a car or a bicycle?" seemed to be inadequate. They reported that capable children often responded incorrectly, apparently understanding "car" to mean a toy car. A pilot test was conducted by adding a different item of similar form in administrations of the MPSI to 210 children. The original item had lower item validity and was replaced by the experimental item: "Which is bigger, a dog or a mouse?"

Acceptable responses on the Information subtest are explicitly specified in the manual. Also, specific questions are provided for the examiner to follow up on certain insufficient, but not inaccurate, responses. For example, a child who holds up fingers to indicate how many eyes he/she has (item 9) is

asked "How many is that?" Several other subtests include such follow-up instructions (Matching, Hopping and Balancing, Counting, and Identifying Body Parts).

Items for Matching subtest were chosen which involve the types of figure discriminations important for letter recognition, but without using actual letters. The first item in this subtest is followed by a demonstration procedure if the child's response is incorrect: (pointing in turn) "See, this looks the same as this one." Similarly, the first items in the Sentence Completion, Prepositions, and Identifying Body Parts subtests are followed by demonstrations if needed.

Administration of the MPSI may be abbreviated for many children by applying the "intermediate decision rule" after item 20 of the Sentence Completion subtest. If the child has a subscore of 8 or less at this point, the test can be discontinued and the child scored as refer. With a subscore of 17 or more, the child can be assigned an outcome of pass at this point.

Hopping and Balancing is the most demanding subtest for the examiner, who must give instructions precisely, demonstrate the tasks adroitly, and observe and score accurately. Wording was greatly simplified to facilitate easy memorization of instructions.

Testing moves quickly through the next four subtests: Naming Colors, Counting, Prepositions, and Identifying Body Parts. In Naming Colors, the child is shown colored blocks in turn and must supply the color name. The Counting subtest requires the child to count out the number of blocks requested, and to count to seven. Identifying Body Parts involves no spoken language; the child points to the body part named.

Repeating Sentences is the last subtest administered. The child, now quite accustomed to the examiner, is more inclined towards verbal production. The five sentences sample a broad range of syntactic elements. The sentences become successively longer and more complex. They are scored for memory and accuracy of word form, not for articulation.

With the speech sample from Repeating Sentences in mind, the examiner rates the child on intelligibility. The child's overall score is then totaled and, referring to the cutoff score for the child's age, an outcome of pass or refer is recorded. More comprehensive psychoeducational assessment is recommended for those children classified as refer.

Research Results

The technical adequacy of the MPSI was evaluated through a number of research studies. Early research data on 728 preschool children verified that the MPSI had high internal consistency. Item-total correlations, as shown in Table 2, were atypically high. Internal consistency reliability for the scale was .95 (Kuder-Richardson formula 20), indicating that the items held together well as part of a single overall scale.

Additional analyses with the data from these 728 children supported the idea of introducing an intermediate decision rule. Examiners had been asking the question of when they should terminate an administration with a child who was performing very poorly and was certain to be referred. Several alternative decision points were tested by comparing the number of items answered correctly up to that point by children who had an ultimate outcome of refer, and those with an outcome of pass. After the first 20 items proved to be the best point for making an early decision to refer (see Table 3). The 73 children who had a score of 8 or less at this point all had a final outcome of refer, therefore

no errors would be made by referring children in this scoring range. Using this rule, the test could have been shortened at this point (probably to the relief of the child and the examiner) for almost half of the 155 referred children in this sample. Coincidentally, this same intermediate decision point was the most efficient point at which to make an early decision of pass. Of children with scores of 17 or more after item 20, all 383 had an outcome of pass after the complete MPSI administration. Following this rule, an early decision could have been made for over two-thirds of the 573 children in this sample ultimately classified as pass.

Reliability of the MPSI was investigated through a combined test-retest/inter-rater reliability study which involved 61 children in eight Minneapolis-St. Paul preschool centers. For this study, nine paraprofessional examiners were trained according to the guidelines provided in the manual. The MPSI was administered twice to 51 children in the sample by two different examiners approximately one week apart. Test-retest reliability was .92 for this group. Reliability figures remained at or above .90 when computed separately for subgroups of younger children (ages 3-8 to 4-4) and older children (4-5 to 5-1). The pass/refer screening outcome was identical on both occasions for all 51 children.

Inter-rater reliability figures were obtained by having an examiner and an observer independently score a child's responses during a live administration. This was done during test administrations of 45 children in the sample. Reliability was computed as the percentage of examiner-observer pairs who agreed on the scoring of an item. Agreement on individual items ranged from 84.4% to 100%. Only two items had less than 90% agreement (item 25--balancing on one foot, 84.4%; and item 50--repeating the sentence "The boy will eat a big breakfast," 86.0%), while over half the items had 100% agreement. Summing

over the items within subtests, inter-rater agreement on subtests ranged from 91.1% (Hopping and Balancing) to 100% (Prepositions). Over the entire set of MPSI items, scoring agreement averaged 97.2%.

Four separate concurrent validity studies, reported in the MPSI manual, show moderate to high correlations between the MPSI and various assessment measures. Of these, I will present the one study which provided results in terms of agreement (i.e., hit rates) rather than correlations. A group of preschool children who were given the MPSI and who were referred from screening on the basis of other criteria were administered the Stanford-Binet Intelligence Scale (Terman & Merrill, 1972) and/or the Preschool Language Scale (Zimmerman, Steiner, & Evatt, 1969) in the course of further assessment. Results in Table 4 show agreement between MPSI outcome and Stanford-Binet scores below 80 (i.e., borderline range and below). All 11 children scoring below 80 on the Stanford-Binet were classified as refer by the MPSI.

The Preschool Language Scale was administered to 48 children in this group. A cutoff score of 85 was used to determine at-risk status on the PLS, since the group mean was five points higher on the PLS than the Stanford-Binet. As shown in Table 5, all except one child scoring below 85 on the PLS received a score of refer on the MPSI.

The most important component of this research program was a large-scale predictive validity study involving follow-up of preschool screening program results. This study covered several lines of research in addition to MPSI validity, such as optimal strategies for applying screening data, and validity of parent reports. A battery of screening procedures, including the MPSI, Developmental Indicators for the Assessment of Learning (DIAL) (Mardell & Goldenberg, 1975), a brief test of speech articulation, and a short

developmental inventory completed by parents was administered to children screened by the Minneapolis Public Schools. The study sample consisted of 428 children, ages 49 to 64 months, who were screened during the school year prior to kindergarten and rated on a follow-up questionnaire by teachers during the spring of their kindergarten year. Since preschool screening took place throughout the school year, the prediction interval from screening to follow-up ratings ranged from one to one-and-a-half years. Return rate for the follow-up rating questionnaires was 97.2%. An additional criterion measure was available for 296 children in the sample who were administered the Metropolitan Readiness Test (Hildreth, Griffiths, & McGauvran, 1964) early in first grade.

Ratings of kindergarten performance were obtained with the follow-up questionnaire shown in Figure 2. Regular and special education teachers rated children in nine areas: learning habits, reading readiness, speech, gross motor, social-emotional development, pre-academic skills, language development, fine motor/perceptual motor, and expected readiness for first grade. Ratings were averaged over the nine categories to classify children into three criterion groups: 29 children (6.8%) with an average rating of under 2.0 were classified as having "moderate to severe problems," 36 children (8.4%) with ratings from 2.0 to 2.5 were classified "mild problems," and 363 children (84.8%) with average ratings over 2.5 were assigned to the "no significant problems" group.

The DIAL was chosen for comparison with the MPSI as it was the most widely used screening instrument in Minnesota at the time. Also, since the MPSI was new and unproven, it was necessary to have an established screening instrument upon which to base referral decisions. Two alterations of standard DIAL procedures were instituted to accommodate screening program

constraints. First, the entire DIAL was administered by the same examiner. Second, children receiving a score of "re-DIAL" were not readministered the areas in which they scored below the cutoff, but were simply referred for further assessment. In reporting of hit rates, these children are grouped with those scored as "refer."

As shown in Table 6, the MPSI had distinctly higher hit rates than the DIAL in predicting kindergarten performance criterion groups. The MPSI correctly referred 25 of the 29 children (86.2%) with moderate to severe problems, and 41 of the 65 children (63.1%) in the two problem categories combined. The DIAL correctly referred 21 of the 29 children (72.4%) with moderate to severe problems, and 35 of the 65 children (53.8%) in the two problem categories combined. In terms of over-referrals (i.e., false positives), the MPSI referred 25 children in the no problem group, compared to 24 over-referrals for the DIAL.

The Metropolitan Readiness Test administered at beginning first grade provided a second criterion measure. Two different criterion groups were established by setting cutoffs at 1 standard deviation and at 1-1/2 standard deviations below the mean according to MRT national norms. In this sample of 296 children, 13 children (4.4%) were over 1-1/2 standard deviations below the mean and 42 children (14.2%) were more than 1 standard deviation below the mean. As shown in Table 7, the MPSI identified 11 of the 13 children in the extreme low group, and 25 of the 42 children who were more than one standard deviation below the mean on the MRT. By comparison, the DIAL identified 8 of 13 children in the extreme low group, and 21 of the 42 children scoring more than one standard deviation below the mean.

Several other findings from this follow-up study are worth noting. First, the speech intelligibility rating subjectively made by examiners had greater validity than a brief (8 item) speech articulation subtest in predicting children with speech problems in kindergarten. The results suggested that the speech intelligibility might best be used as a "pre-screening" measure, with only those children receiving the highest rating of 4 excluded from further screening. Following this rule, the 23.7% of the children in this sample with ratings of 1 through 3 would have been screened further for speech.

Another finding was that screening decisions based on the combined results of MPSI, DIAL, and the parent developmental inventory yielded little improvement in predictions over the MPSI alone. The language and self-help scales of the developmental inventory showed promise by adding significant variance to a regression equation, but resulted in only slight hit rate gains. Still, given the short length of the developmental inventory (especially the language scale of just four items), these may be regarded as favorable findings on the prospective value of parent report information.

The question of how far into the future screening results may be applied was also addressed. Other studies have shown that prediction of later academic performance increases to substantial levels by late in the kindergarten year (Eaves, Kendall, & Crichton, 1974; Novack, Bonaventura, & Merenda, 1973). Similar results were found in this study. The kindergarten teacher ratings were superior to the screening battery in predicting performance on the Metropolitan Readiness Test in first grade. Clearly, screening data decreases in value over time, and should be superseded by information on the child's actual functioning in the educational setting. Further research hopefully will indicate at what point in the kindergarten year teacher ratings become more valid than preschool assessment.

Reflections and Recommendations

Returning to the hit rate results comparing the validity of MPSI and DIAL decisions (Tables 5 and 6), these findings surpassed my most optimistic expectations for the MPSI. I had waited long and patiently before advising the Minneapolis Public Schools that this was a viable screening instrument and worthy of being published. However, by the time these results came in, my philosophical position had shifted in a direction that led me to endorse this screening instrument, as well as the practice of preschool screening, with considerable caution.

A critical finding of the follow-up study was that the MPSI, despite its clear edge over the DIAL in this particular situation, still failed to identify some children who were rated as having moderate to severe problems in kindergarten, and many of those children with only mild problems. Further research in the form of a comprehensive survey of early identification follow-up studies (Lichtenstein, 1980b) revealed that this is invariably the case. In fact, the seemingly humble hit rates obtained here proved to be higher than in any other study involving unbiased early prediction of school problems. (Note: "unbiased," meaning the screening or identification decisions were established in advance; "early," meaning preschool or kindergarten level; "prediction," meaning six-month or longer follow-up.) It was also found that there was no consistent relationship between length/comprehensiveness of identification procedures and their accuracy. It appears that, in direct assessment of the preschool child, the point of diminishing returns is fairly quickly reached.

There are two main explanations that may account for these findings. The first is that the state of the art of early identification is "developmentally delayed"--a criticism often found in the literature--and that

technological advances are necessary to bring about much-need improvements in accuracy.

The second explantion is, to my mind, more significant and to the point. Errors of identification are simply unavoidable in early identification. This is due to several reasons. Reliability of measures is limited by the young child's variability in mood, in motivation, in attention span, and in adapting to novel situations. Imprecise definitions and variations in measurement procedures result in inconsistencies in classifying children into categories of "at risk" or "not at risk." The very nature of associating the developmental status of children with discrete categories leads to errors. Children near the borderline between categories tend to alternate between "at risk" and "not at risk" status depending on the situation and on the criterion used. However, the greatest limitation upon trying to predict the future status of young children is imposed by measurement stability--or rather, the lack of it. Preschool children are likely to make dramatic changes over time. The child who is accurately identified as being "at risk" may show spontaneous growth that makes him/her indistinguishable from peers two years later.

These factors impose such a strong restrictive impact upon psychoeducational predictions made about preschool children that we should simply accept our modest success to date in early identification as an accurate reflection of the nature of the task. The time has come to stop looking for more elaborate measures to bring about a breakthrough in early assessment and to draw some conclusions about the limits of early identification accuracy. While there are still technical improvements to be made in preschool screening procedures, by and large

> "the precision of the measurement procedure can never compensate for the uncertainties of trying to definitively classify a changing organism over time, and cannot surpass certain limits on prediction rates." (Lichtenstein, 1980b)

We must come to acknowledge that it is not possible to eliminate identification errors, and that we should pursue early identification practices that deal realistically with this limitation. We need to reverse the current trend of investing excessive resources in trying to definitively identify and diagnose those preschool children in the "borderline" area (i.e., with mild problems), a task which borders on the impossible. It is counterproductive to use elaborate and expensive identification procedures to assess preschool children at a single point in time (although, we are encouraged to do so by P.L. 94-142). Early identification programs should be designed to use their resources wisely, and to minimize the negative consequences of identification errors. To this end, the following recommendations are proposed:

1. Use brief, but technically adequate, procedures for screening preschool children.
2. Supplement data from standardized testing with other cost-effective sources of information, e.g., parent reports, ratings from preschool teachers, and behavioral observations during screening.
3. Screen preschool children periodically, particularly children in high risk populations. Also, children close to the cutoff point who are not referred should be rescreened after a shorter interval than usual.
4. Conduct follow-up assessment only to the extent needed. Difficult identification decisions may be best resolved by further assessment at a later point. Children placed in intervention

programs can be assessed in an ongoing manner within the context of the program.

5. Do not rely upon preschool screening results beyond the point when valid information about the child's actual functioning in an educational setting is readily available.

One final, though perhaps unnecessary, caveat . . . I have presented the MPSI to you as a simple, sensible, and effective screening instrument. The research conducted so far bears this out. But this must be tempered by the fact that current normative and validity data are limited to a particular geographic area. I am eager to see the MPSI used in other screening situations, and hope that this will provide opportunities for more research to be conducted. At a minimum, other screening programs using the MPSI should devise local norms which correspond to the needs of the specific screening situation. I invite interested parties to consult with me in planning screening research projects, and to share the results they obtain with the MPSI.

REFERENCES

Davidson, J., Lichtenstein, R., Canter, A., & Cronin, J. Directory of developmental screenings instruments. Minneapolis: Minneapolis Public Schools, 1977.

Eaves, L. C., Kendall, D. C., & Crichton, J. U. The early identification of learning disabilities: A follow-up study. Journal of Learning Disabilities, 1974, I, 632-638.

Hildreth, G., Griffiths, H., & McGauvran, M. Metropolitan Readiness Test. New York: Harcourt, Brace & World, 1964.

Kaufman, A. S. Factor structure of the McCarthy Scales at five age levels between 2-1/2 and 8-1/2. Educational and Psychological Measurement, 1975, 35, 641-656.

Lichtenstein, R. Minneapolis Preschool Screening Instrument. Minneapolis: Minneapolis Public Schools, 1980a.

Lichtenstein, R. Early identification of psychoeducational problems: Assessing the limits of prediction. Paper presented at the Third International Conference on Early Identification of At Risk Children, Teton Village, Wyoming, September 1980b.

Mardell, C. D., & Goldenberg, D. S. DIAL - Developmental Indicators for the Assessment of Learning: Manual. Highland Park, IL.: DIAL, Inc., 1975.

Meyers, C. E., Orpet, R. E., Attwell, A. A., & Dingman, A. F. Primary abilities at mental age six. Monographs of the Society for Research in Child Development, 1962, 27, No. 1.

Novack, H., Boneventura, E., & Merenda, P. A scale for early detection of children with learning problems. Exceptional Children, 1973, 40, 98-105.

Ramsey, P. H., & Vane, J. R. A factor analytic study of the Stanford-Binet with young children. Journal of School Psychology, 1970, 8, 278-284.

Terman, L., & Merrill, M. Stanford-Binet Intelligance Scale, 1972 norms edition. Boston: Houghton Mifflin, 1973.

Wallbrown, F. H., Blaha, J., & Wherry, R. J. The hierarchical factor structure of the Wechsler Preschool and Primary Scale of Intelligence. Journal of Consulting and Clinical Psychology, 1973, 41, 356-362.

Zimmerman, I. L., Steiner, V. G., & Evatt, R. L. Preschool Language Scale. Columbus, Ohio: Charles E. Merrill, 1969.

Table 1

MPSI Normative Data

Age Range	N	Mean	Standard Deviation	Standard Error of Measurement	Passing Score
3-7 to 3-10	124	35.5	14.2	4.1	25
3-11 to 4-1	141	36.3	14.6	4.2	28
4-2 to 4-4	250	38.8	12.4	3.6	30
4-5 to 4-7	270	41.2	11.4	3.3	33
4-8 to 4-10	251	42.9	11.9	3.4	35
4-11 to 5-1	206	43.2	12.6	3.6	37
5-2 to 5-4	78	44.6	9.3	2.7	38
Total Sample	1,320	40.5	12.7	3.7	

Table 2

Minneapolis Preschool Screening Instrument

Item Analysis Data

Subtest	Item No.	Point Biserial Correlation with MPSI Total	Phi Correlation with MPSI Decision	Proportion Passing
Building	1	.33	.27	.96
	2	.43	.37	.87
	3	.47	.41	.82
	4	.42	.31	.54
Copying Shapes	5	.41	.34	.84
	6	.56	.46	.70
	7	.55	.47	.64
Information	8	.35	.24	.95
	9	.58	.50	.86
	10	.69	.63	.79
	11	.58	.52	.65
	12[a]	.52	.38	.89
	13	.63	.55	.61
Matching	14	.50	.39	.91
	15	.51	.45	.84
	16	.47	.39	.79
	17	.44	.37	.68
Sentence Completion	18	.53	.47	.72
	19	.51	.45	.82
	20	.62	.56	.66
	21	.60	.55	.64
	22	.51	.46	.60
Hopping and Balancing	23	.45	.33	.85
	24	.44	.31	.77
	25	.39	.29	.69
	26	.40	.30	.58
Colors	27	.53	.48	.80
	28	.57	.49	.80
	29	.64	.60	.77
	30	.53	.42	.77
	31	.61	.54	.77
Counting	32	.63	.53	.84
	33	.64	.49	.65
	34	.66	.55	.70
	35	.55.	.45	.90
	36	.64	.53	.70
Prepositions	37	.43	.34	.96
	38	.46	.35	.94
	39	.58	.49	.88
	40[b]	.59	.50	.83
Body Parts	41	.42	.30	.95
	42	.51	.46	.91
	43	.45	.36	.81
	44	.55	.46	.84
	45	.41	.33	.42
Repeating Sentences	46[c]	.41	.32	.95
	47	.58	.50	.89
	48	.44	.32	.84
	49	.58	.50	.81
	50	.54	.46	.73

Note. N=728

[a] Based on N=210

[b] Based on N=105

[c] Based on N=608

Table 3

Intermediate Decision Rule: Subscore Distribution by Final MPSI Decision

		Subscore after Item 20			
		0-8	9-16	17-20	
Final MPSI Decision	Pass	0	190	383	573
	Refer	73	82	0	155
		73	272	383	728

TABLE 4

CLASSIFICATIONAL AGREEMENT BETWEEN MPSI OUTCOME

AND STANFORD-BINET INTELLIGENCE SCALE

	STANFORD-BINET		
	BELOW 80	80 AND ABOVE	TOTAL
REFER	11	12	23
PASS	0	12	12
TOTAL	11	24	35

TABLE 5

CLASSIFICATIONAL AGREEMENT BETWEEN MPSI OUTCOME

AND PRESCHOOL LANGUAGE SCALE

PRESCHOOL LANGUAGE SCALE

	BELOW 85	85 AND ABOVE	TOTAL
REFER	8	22	30
PASS	1	17	18
TOTAL	9	39	48

Table 6

Classificational Agreement Between MPSI and DIAL Outcome and Teacher Rating Criterion Group

Overall Teacher Rating

		Severe-to-Moderate prob.	Mild problem	No problem		
DIAL Outcome	Refer	20	10	17	47	(11.0%)
	Re-DIAL	1	4	7	12	(2.8%)
	OK	8	22	339	369	(86.2%)
		29	36	363	428	

Overall Teacher Rating

		Severe-to-Moderate prob.	Mild problem	No problem		
MPSI Outcome	Refer	25	16	25	66	(15.4%)
	Pass	4	20	338	362	(84.6%)
		29	36	363	428	

Table 7

Classificational Agreement between MPSI, DIAL Outcomes and First Grade Metropolitan Readiness Test Scores

MRT Score (in standard deviations)

DIAL Outcome	<-1.5	-1.5 to -1.0	>-1.0	Total
Refer	5	11	11	27
Re-DIAL	3	2	4	9
OK	5	16	239	260
	13	29	254	296

MPSI Outcome	<-1.5	-1.5 to -1.0	>-1.0	Total
Refer	11	14	19	44
Pass	2	15	235	252
	13	29	254	296

NAME _____ (last) _____ (first)

ADDRESS _____ _____

ID#_____ SEX: MALE FEMALE

DATE _____ (year month day) SCREENING LOCATION _____

BIRTH-DATE _____ (year month day) EXAMINER _____

AGE _____ (year month day)

PASS ☐

REFER ☐

CUT-OFF SCORES

AGE	PASSING SCORE
3:7-3:10	25
3:11-4:1	28
4:2-4:4	30
4:5-4:7	33
4:8-4:10	35
4:11-5:1	37
5:2-5:4	38

COMMENTS

BUILDING

1.___
2.___
3.___
4.___

COPYING SHAPES

5.___
6.___
7.___

INFORMATION

8.___ What do you wear on your feet?
9.___ How many eyes do you have?
10.___ What animal barks?
11.___ What is a window made of?
12.___ Which is bigger, a dog or a mouse?
13.___ What do we do with our ears?

MATCHING

14.___
15.___
16.___
17.___

SENTENCE COMPLETION

18.___ Brother is a boy, sister is a ----.
19.___ Planes go up, rain comes ----.
20.___ A block is square, a ball is -----.
21.___ Running is fast, walking is ----.
22.___ A bird flies, a fish -----.

HOPPING BALANCING

23.___ Hops 1 time — 1st foot
24.___ Hops 1 time — 2nd foot
25.___ Balance 5 sec. — 1st foot
26.___ Balance 5 sec. — 2nd foot

COLORS

27.___ Green
28.___ Orange
29.___ Black
30.___ Yellow
31.___ White

COUNTING

32.___ 2 Blocks
33.___ 4 Blocks
34.___ 3 Blocks
35.___ Counts to 3
36.___ Counts to 7

PREPOSITIONS

37.___ On
38.___ Inside
39.___ Under
40.___ Beside

BODY PARTS

41.___ Knee
42.___ Thumb
43.___ Chin
44.___ Elbow
45.___ Ankle

REPEATING SENTENCES

46.___ (The) girl is big.
47.___ Sit on the kitchen floor.
48.___ The dog won't come home.
49.___ They like to play in the rain.
50.___ The boy will eat a big breakfast.

_____ **TOTAL SCORE**

SPEECH INTELLIGIBILITY

___ 4-Always understandable
___ 3-Most words understandable
___ 2-Some words understandable
___ 1-Few words understandable

30

MINNEAPOLIS PRESCHOOL SCREENING INSTRUMENT

31

Copyright 1980, Minneapolis Public Schools
(reprinted with permission)

Figure 1
MPSI Scoring Form

Figure 2

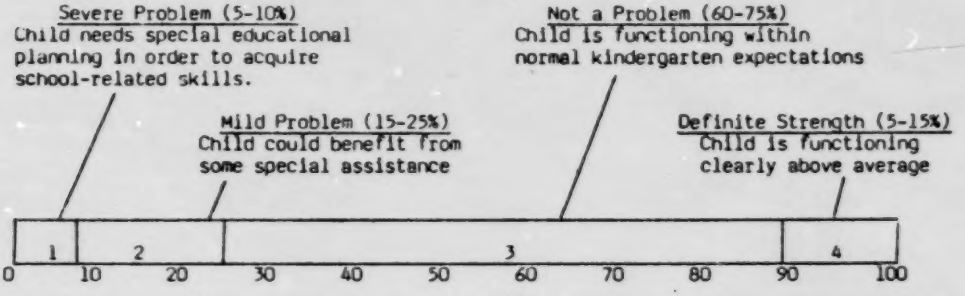
Follow-up Questionnaire

Child's Name _____ (Last) _____ (First) TR _____

Birthdate _____ (Month) _____ (Day) _____ (Year) Sex _____ (male=1, female=2)

Rater's Name and Position _____ School _____

A. Please rate the child's performance on a scale of 1 to 4 for each of the areas described. The chart below explains each of the four ratings and shows the approximate percentages of kindergarten children that would be expected to fall into each category. Note that a child is most likely to receive a rating of 3.



Kindergarten Performance Area	Severe Problem	Mild Problem	Not a Problem	Definite Strength
Learning Habits (Paying attention, following directions)	1	2	3	4
Reading Readiness (Visual and auditory discrimination, letter knowledge)	1	2	3	4
Speech (Articulation, voice quality)	1	2	3	4
Gross Motor (Walking, running, hopping, balancing)	1	2	3	4
Social-Emotional Development (Relationships with children and adults, frustration tolerance, emotional stability)	1	2	3	4
Pre-Academic Skills (Comprehension, memory, learning skills, concept knowledge, reasoning)	1	2	3	4
Language Development (Understanding language, vocabulary expressing ideas)	1	2	3	4
Fine Motor/Perceptual Motor (Copying, cutting, drawing, pasting)	1	2	3	4
Expected Readiness for First Grade (Based on current school performance)	1	2	3	4

Copyright 1980, Robert Lichtenstein